2016 年数据新闻经典案例和发展趋势

文 | 戴　玉　李佳璇　操子宜

2016 年是中国数据新闻如火如荼发展的一年，无论是传统纸刊还是门户网站，无论是市场化媒体还是体制内的央媒，无论是业界还是高校，都开始涉足数据新闻报道，而且也都拿出了让人印象深刻的数据新闻作品。

2016 年数据新闻有了很明显的“进阶”趋势，数据库新闻成为最大亮点，各媒体都纷纷建立自己独家的“新闻数据库”，搭载网页互动技术来开放查询、匹配功能。以往的大部分数据新闻多以数据结论、统计结果作为报道对象，民众接触到的主要是经过记者挑选、专家解读过后的结果，无法接触到最初始的数据源，更无法对其进行查询、分析。

这主要是有三个原因，一是过往的平面媒体互动性差，无法在有限的版面内呈现大规模数据。二是受众对数据的解读能力有限，媒体看不到把数据库进行公开的必要性。三是媒体建立、报道数据库的能力和对数据库的关注度都不够。

但是，这种情况出现了集中转变，比较有代表性的是凤凰资讯在魏则西事件愈演愈烈之时推出的 H5 能够查询“你居住的城市有多少莆田系医院”。当大家把矛头指向莆田系医院的时候，这则作品通过数据库新闻非常及时地解决了大家最关心的问题，而且查询的方式也让每个人都可以从数据库中抓出“自己身边的新闻”，引起广泛传播。

同样用查询类数据库 H5 来引爆朋友圈的还有腾讯新闻“事实派”栏目推出的《纪委你好，干得漂亮——十八大以来全国落马官员案例汇总》。这个涵盖了 2012 年 11 月 8 日到 2015 年 12 月 31 日纪检监察机关、司法部门等公开通报的各省科级及以上落马官员资料的数据库，囊括了约 19000 多名落马官员，所提供的功能同样是可以查询任意城区、城市、省份的落马官员的名字、级别、所犯罪名等。除了查询之外，该 H5 也提供了一定的统计展示功能，比如可以看到各省官员落马的罪名原因统计情况，各省落马官员总人数对比等等。

数据库新闻的特点就是数据量大、案例全，几乎每个读者都能根据需要去查询出属于自己或本地的结果，也没有特定的主要新闻结论的指向，实现了一种“千人千面”式的新闻报道。然而另一个角度来看，这种数据新闻耗时比较长，再加上制作和设计周期，整个作品留给采访解读、挖掘分析的时间就不够多。

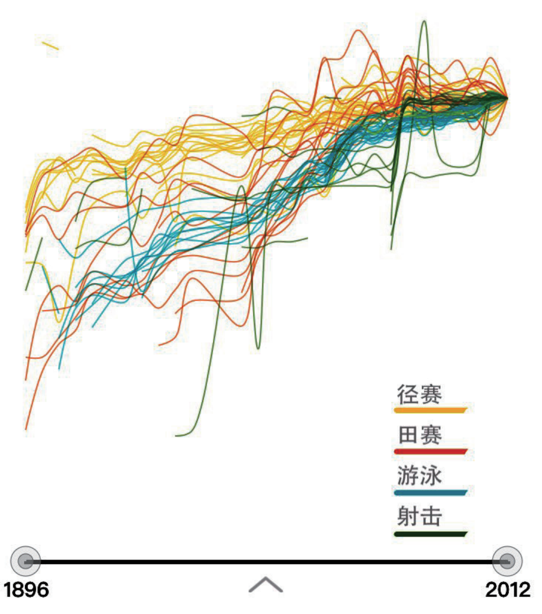
除了查询类，2016 年也出现了一种匹配型的游戏类数据库新闻 H5，这就是中国网和图政数据联合推出的对全国人大代表的组合报道。这组报道包括两个数据库，一个是年轻人对全国人大代表看法和观念的网络民意调查结果，一个是全国人大代表数据库。在后一个数据库中，包含了通过公开新闻，手工逐一搜索出的代表们的现任职务信息，以及相匹配的行业。

《谁能代表我？》这一 H5 便是基于代表们的数据库，揣摩受众对代表性的理解，设计出了一套匹配算法。每个读者输入自己的性别、年龄、民族、行业、党派、所在省份信息，就可以自动计算匹配度并呈现出最匹配的人大代表的信息。

实际上，数据库新闻的长期开发应用价值应远远高于一条即时新闻。经过漫长收集过程而积淀下来的数据库，可以用网页的形式永久公开和定时更新。但遗憾的是，目前对新闻中所使用的数据库的保存、后期挖掘和结构化公开都不太到位。长此以往，这种数据库新闻可能很难获得持续的投入支持，可能也很难进一步提高挖掘深度和数据质量。

每一根曲线由一个项目上历届奥运会的冠军成绩相连而成。纵坐标表示该名选手在其比赛项目上的长度、高度或速度成绩，位置越高则成绩越好。其中，跑步及游泳等项目的时间成绩换算成了速度成绩，以模拟各届冠军同场比拼的结果。





不过，随着数据库新闻热潮而来的，是 2016 年中国数据新闻在数据可视化方向上的明显精进。人民日报中央厨房《1976–2016 地震全纪录》H5 通过录屏的方式，实现了百万级的全球地震数据在移动端的立体化展示，并可以让受众互动查看近 40 年中国地震的分布情况。

不过真正在数据可视化的表达设计和互动上有突破的，还是当推《第一财经周刊》新一线城市研究所的《你真的了解你所在的城市吗？》竞猜类 H5。这个作品包含了 300 多个城市的衣食住行各类数据，比如在优酷上的人均收看视频时长、地铁运营时间、一顿饭 / 一场电影的人均消费额，以及餐厅、民用机场、宠物店、超高层建筑、咖啡馆等场所的数量，受众可以去猜测具体数值。

而在猜测数值的互动方式上，该 H5 就突破了常见的柱状或饼状图的图表形式，也丰富了简单的窄块状点选型和直线拖动型互动的呈现效果。比如受众可以依照圆弧形去拨动钟表盘、仪表盘、圆周，或者逐一上下拨动 4 列“算盘珠”式样的数值来进行竞猜。这些形式在以往的数据新闻互动动作中都是比较少见的。

此外，南京艺术学院设计学院的视觉信息设计作品也在 2016 年大放异彩。在全球数据可视化高端赛事“信息之美”大赛中，中国赛区共计 18 件作品入围，南京艺术学院以 6 件入围作品位列国内第一，它们分别是：《垃圾分类回收图表》《图解现代手工艺——陶瓷工艺》《记忆 Memory》《二十四节气信息图表》《伽利略》《一个机场是如何运转？》。

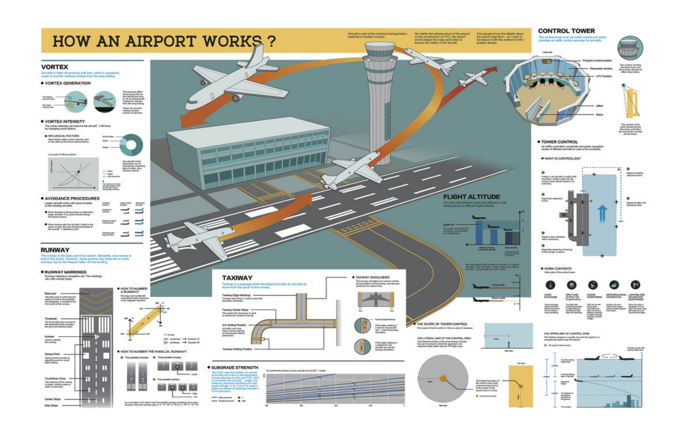
“学院派”的数据新闻作品以平面设计为主，在新闻性上稍弱，除了数字之外也有很多以展示结构、历史、过程为主的作品。但相对的，他们也有更多时间去专注于梳理某一个命题，并倾注到设计感之上。

除了走向“数据更全”和“设计更美”之外，中国数据新闻在 2016 年也在努力靠近“更新更快”的目标。凤凰资讯的《美国大选开票直击：实时票数统计》就在美国大选最终投票当天呈现“刷屏级”的效果，受众通过手机就能随时查看两位候选人的实时得票结果与对比。该 H5 数据准确而且系统稳定、设计简约，在中国的数据直播上可以说是有一定开创性的。也可以看到，凤凰资讯的数据新闻作品一般来说新闻性非常强，似乎并未刻意强调数据新闻的数据属性，但却将数据新闻的新闻时效性运用得较为突出。

就如香港《南华早报》在 2016 年获得“信息之美”设计大赛社区奖（由全球数据可视化社区通过网上投票得出）金奖的作品《1996–2015 中国各省经济增长》，在数据上未见得能令中国人感到新颖，但却在国际投票中让外国人对中国的增速有了直观感受，真正有了新闻的价值。

然而无论数据新闻如何发展，最终还是要走向纵深解剖的综合报道，将数据库、新闻点（数据新发现）、采访、可视化等都融合到一起。

DT 财经《10 万条挂号大数据显示：互联网也救不了看病难》的报道就是从挂号的数据中来分析哪些省份的医院更支持网上预约、哪些医院的预约量“过载”、哪些医院



评分不高等。澎湃新闻 2016 年推出的《奥运人类突破曲线：极限在哪里》在基于奥运极限值的极简曲线之上，加了精练的点评，在数据结论和图表的配合上更进了一步。央视 2016 年推出的《中国舆论场》和“中国舆论场指数”，就是基于大数据分析全媒体平台的舆论热点，并请来相关专家对热点话题进行解读。

可以说，在扭转了“重解读、轻数据”的传统观念之后，2016 年的中国数据新闻界迈向了“数据为王”的数据库新闻的领地，虽然仍然努力在数据分析、解读和采访方面进行突破，然而仍旧少有融合多方佐证的数据新闻深度调查报道出现，甚至在某种程度上保留了数据解读的弱化和数据采访的相对缺失。

数据新闻毕竟是一种分工众多、流程繁杂的新闻种类，涉及数据收集、数据分析、采访解读、报道或互动文案写作、可视化、技术支持等多方面内容。在发展早期，很可能会偏重某处或忽略某方面。

然而，若论作品的综合水平，还是当推财新网的《从调控到刺激：楼市十年轮回》。从新闻角度来看，这一作品文字冷峻、成熟，涵盖人物案例采访、数据分析和政策研究，新闻各个内容部分相对完善。即便是在数据可视化方面，也达到比较高的水准。尝试了数据的 3D 互动展现，用棋盘图和曲线图联动的形式呈现了 70 个城市 2005 年 7 月到 2016 年 7 月的房价涨跌幅度。互动设计则贴近受众需求，只需要输入城市名称和家庭月收入，就能自动计算在该地买一套 80 平米的房子需要不吃不喝多少年。哪怕是对于数据来源及其计算方法的标注，也是详尽而规范。

从整体来看，中国数据新闻在 2016 年确实在数据、新闻、可视化三个主要方向皆有推进，而且在移动端发力比较明显。然而平心而论，这种推进只能说是预料之内和情理之中的，在保持过往水准的基础上做出了一定的努力和改进，却尚未达到爆发式成长和出现明显突破的地步。

如果要预测的话，2017 年，数据新闻可能在可视化方面会有更多进展，数据新闻的综合性深度报道不一定能有更多挖掘。相信未来在国际的新闻舞台上，也会有更多中国的数据新闻作品出现。媒

（**作者简介**：戴玉，《南风窗》原编辑，人民日报原中央厨房数据新闻项目负责人，率团队获得中国首届数据新闻大赛二等奖，入围 2016 英国信息之美设计大赛长名单）